AD-A243 357

DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

ank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	Summer 1991	THESIS/DISSERTATION

4. TITLE AND SUBTITLE
Image Processing Applications for Nonlinear Decomposition and Synthesis

5. FUNDING NUMBERS

6. AUTHOR(S)
Eric W. Kelm, 2d Lt

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
AFIT Student Attending: University of Washington

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/CI/CIA-91-065

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFIT/CI
Wright-Patterson AFB OH 45433-6583

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release IAW 190-1
Distributed Unlimited
ERNEST A. HAYGOOD, Captain, USAF
Executive Officer

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

14. SUBJECT TERMS

15. NUMBER OF PAGES
25

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

Image Processing Applications for Nonlinear Decomposition and Synthesis

by

Eric W. Kelm, 2Lt., U.S.A.F.

Master of Science in Electrical Engineering (M.S.E.E.)
University of Washington
Summer, 1991

25 pages

ABSTRACT

Data representation, which is often overlooked in many image processing and analysis applications, is as critical as the algorithms applied to that data. When data is represented properly, simple algorithms can be much more powerful than sophisticated and often complex algorithms applied to an improper representation. In an image, the useful information is generally mixed in with irrelevant information or noise and often it is difficult for the computer to separate the useful information from the large volume of irrelevant data without destroying much of the useful data. To provide a solid foundation for a good solution to this problem, "multiresolution decomposition and synthesis" approaches have been developed which decompose raw image data into a set of partial information channels, where each channel represents a certain modality (or aspect) of the raw image. The channels can then be processed individually or cooperatively with a wide variety of results possible. After processing of the channels is completed, they can be selectively synthesized to recover the original image, an improved version, or an image with certain features highlighted. In this way, the irrelevant information can be more effectively rejected or ignored while the useful data can be used or modified in whichever manner is desired. For this project, I have studied the usefulness of the nonlinear image pyramid, compared to the linear type of pyramid and raw data format, in the areas of image compression, image transforms, and image enhancement via noise removal.

References

[1] H. C. Andrews and C. L. Patterson, "Single Value Decomposition (SVD) Image Coding." *IEEE Transactions on Communications*, April 1976.

[2] H. C. Andrews and C. L. Patterson, "Single Value Decompositions and Digital Image Processing." *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 24: 26 - 36, February 1976.

[3] P. J. Burt and E. Adelson. "The Laplacian Pyramid as Compact Image Code." *IEEE Transactions on Communication*, 31: 532 - 540, 1983.

[4] A. K. Jain. ***Fundamentals of Digital Image Processing***. Prentice-Hall: New Jersey, 1990.

[5] E. W. Kelm, J. N. Hwang, J. S. J. Lee, "Image Enhancement via Multiresolution Decomposition and Synthesis." Submitted to the *1992 ICASSP Proceedings*, San Francisco, March 1992.

[6] J. S. J. Lee, J. N. Hwang, and E. W. Kelm, "Image Analysis by Information Decomposition and Synthesis." To appear in the *25th Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, November 1991.

[7] Y. Lu am R. C. Vogt, "Multiscale Analysis Based on Mathematical Morphology." *SPIE Proceedings on Image Algebra and Morphological Image Processing II*, July 1991.

[8] S. Ranganath, "Image Filtering Using Multiresolution Representations." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13: 426 - 439, May 1991.

[9] P. Salembier and L. Jaqueoud, "Adaptive Morphological Multiresolution Decomposition." *SPIE Proceedings on Image Algebra and Morphological Image Processing II*, July 1991.

[10] T.H. Yu and S.K. Mitra, "A Simple Image Analysis / Synthesis Technique and Its Application in Image Coding." *SPIE Image Processing Algorithms and Techniques*, 1244: 161 - 170, 1990.

Image Processing Applications for Nonlinear Decomposition and Synthesis

by
Eric W. Kelm

for
Professor Jenq-Neng Hwang
and
Professor Linda Shapiro

E.E. 600A
Department of Electrical Engineering
University of Washington
Summer, 1991

ABSTRACT

Data representation, which is often overlooked in many image processing and analysis applications, is as critical as the algorithms applied to that data. When data is represented properly, simple algorithms can be much more powerful than sophisticated and often complex algorithms applied to an improper representation. In an image, the useful information is generally mixed in with irrelevant information or noise and often it is difficult for the computer to separate the useful information from the large volume of irrelevant data without destroying much of the useful data. To provide a solid foundation for a good solution to this problem, "multiresolution decomposition and synthesis" approaches have been developed which decompose raw image data into a set of partial information channels, where each channel represents a certain modality (or aspect) of the raw image. The channels can then be processed individually or cooperatively with a wide variety of results possible. After processing of the channels is completed, they can be selectively synthesized to recover the original image, an improved version, or an image with certain features highlighted. In this way, the irrelevant information can be more effectively rejected or ignored while the useful data can be used or modified in whichever manner is desired. For this project, I have studied the usefulness of the nonlinear image pyramid, compared to the linear type of pyramid and raw data format, in the areas of image compression, image transforms, and image enhancement via noise removal.

1.0 Introduction

Multiresolution decomposition techniques are commonly used in variety of signal processing applications [Salembier 26], including but not limited to image compression, progressive transmission, enhancement, restoration, analysis, et c.[Ranganath 426]. The classical approach to image decomposition has been the repeated applications of a linear, usually Gaussian, type of lowpass filter to obtain a frequency sensitive decomposition. However, in many cases this may not be the best choice as some applications are better suited for a size or shape dependant decomposition. Thus, instead of a decomposition resulting in a division of frequency space, it may be desirable to have a decomposition of the pattern spectrum.

When our research began, our aim was to investigate the applicability of morphological decompositions to the area of image compression. It is well known that the linear type of pyramid can yield compact image code; could the nonlinear decomposition lead to comparable or even better general compression results?

Image compression can also be achieved through the use of image transforms, such as the Discrete Cosine transform, K-L transform, Single Value Decomposition, et c. Transform

coding, also known as block quantization, involves a unitary transformation of each block so that a large fraction of the total energy is packed into relatively few transform coefficients. The image can then be reconstructed from these few coefficients with a minimal amount of error. The next step was to examine the possibility that the image pyramids could be more effectively transformed than the transformation of the original image with respect to image compression.

Finally, the linear image decompositions can also be used for image enhancement and restoration. Restoration of noisy images and undesirable artifact removal are just two of the possible areas that the pyramid structure can be effectively used. Also, because of the multiscale nature of the pyramids, some types of filtering can be done much more efficiently on the pyramid planes than on the raw image data [Ranganath 437]. In this area, the possibility that morphological pyramids could outperform the linear pyramids and equivalent types of operations on raw image data was investigated.

2.0 General Pyramid Concept

The pyramid image decomposition technique is an iterative procedure designed to isolate image features at different scales and to support more efficient scaled neighborhood operations through reduced image representation [Lee 4]. In this section, an overview of how the pyramids are constructed and then how the image can be synthesized from the data will be given. Possible types of pyramid processing are discussed in great detail in a later section. For the general explanation of the pyramid structure, terminology for linear pyramids will be used as it is more widely understood.

2.1 General pyramid Structure

First, a lowpass decomposition of the original image is accomplished. The zeroth level of the lowpass pyramid will be the original image, **I**. The image is then (linearly or morphologically) filtered and down-sampled to obtain the next level. This is applied repetitively to obtain the entire lowpass pyramid.

$$L_0 = I$$

$$(L_i = D(F(L_{i-1}))), \text{i=>1}$$

'D(.)' represents the down-sizing operation which is simply a decimation by two in each dimension, so for a two dimensional image, the down-sized image has 1/4 the pixels of its parent. 'F(.)' is the filtering operation, usually convolution for the linear case and some

type of morphological operation (Erosion, Dilation, Closing, Opening, et.c.) for the non-linear application.

The bandpass pyramid is then generated using the lowpass pyramid by simply subtracting each lowpass decomposition by the next lower level in the lowpass pyramid. However, the lower level differs in size by a factor of 2 in each dimension. So, the image must be up-sized, U(.), which is a point replication operation, and then interpolated, P(.). The interpolator can be of whatever type the user desires. Often for linear pyramids, a Gaussian type of lowpass filter is used, while the same or a morphological interpolator can be used for the nonlinear pyramids. Thus the bandpass decomposition, Bi, can be specified in terms of the lowpass decompositions. These bandpass images are often referred to as prediction error or residuals because they are not true bandpass images due to the errors introduced by the non-ideal filters and interpolator. Figure 1 graphically shows the construction pro-

$$B_i = L_i - P\,(U\,(L_{i+1}))$$

cedures for both the lowpass and bandpass pyramids.

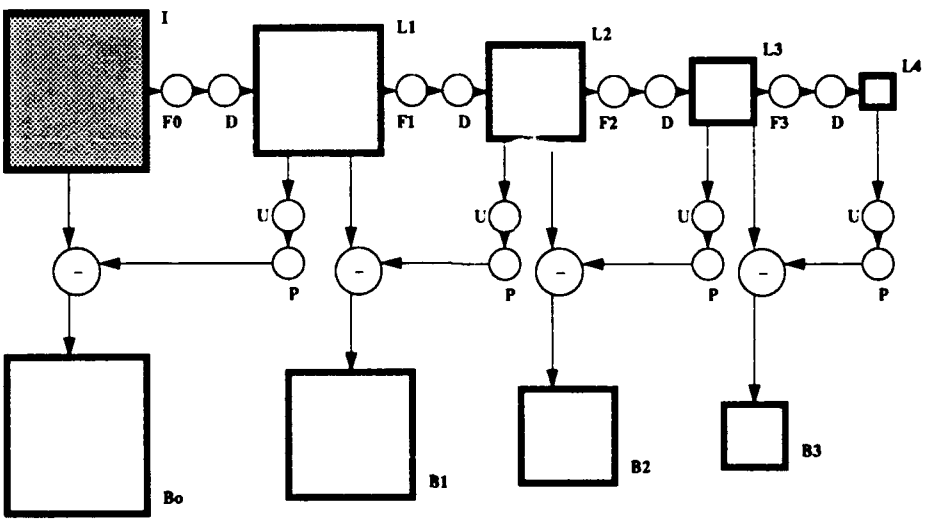

FIGURE 1. Construction procedures for the lowpass and bandpass decompositions

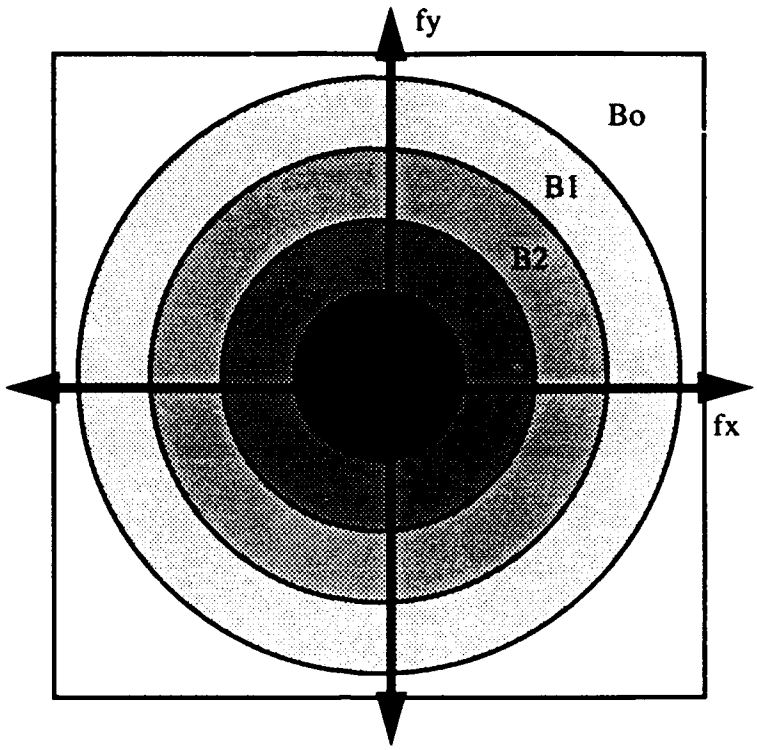

FIGURE 2. 2-D Spatial Frequency / Pattern Spectra Plot of Bandpass Pyramid

Figure 2 shows, graphically, how the bandpass pyramid will ideally look as the bandpass decomposition of an image yields different bands of either spatial frequency or morphological pattern spectrum [Kelnı 3]. It must be noted that due to errors introduced by interpolation and less than ideal filters the actual frequency/pattern spectra plot will not be this clean. One large advantage the pyramid structure has over some other types of representations, such as a Fourier transform, is that the bandpass images retain their spatial locality. Shapes of interest or even single objects can be enhanced without modifying the rest of the image.

One important property of this type of image representation is that the image is represented completely. Full reconstruction is possible without using all of both pyramids. As shown in Figure 1, an order 'k' pyramid will have k distinct levels, L_0 --> L_k , in the lowpass pyramid and k levels, B_0 --> B_{k-1}, in the bandpass pyramid. However, only the smallest lowpass image, L_k , and the bandpass pyramid are needed for full image reconstruction. It will later be proven that at most 4/3 of the pixels in the original image are required to store the entire bandpass pyramid and L_k .

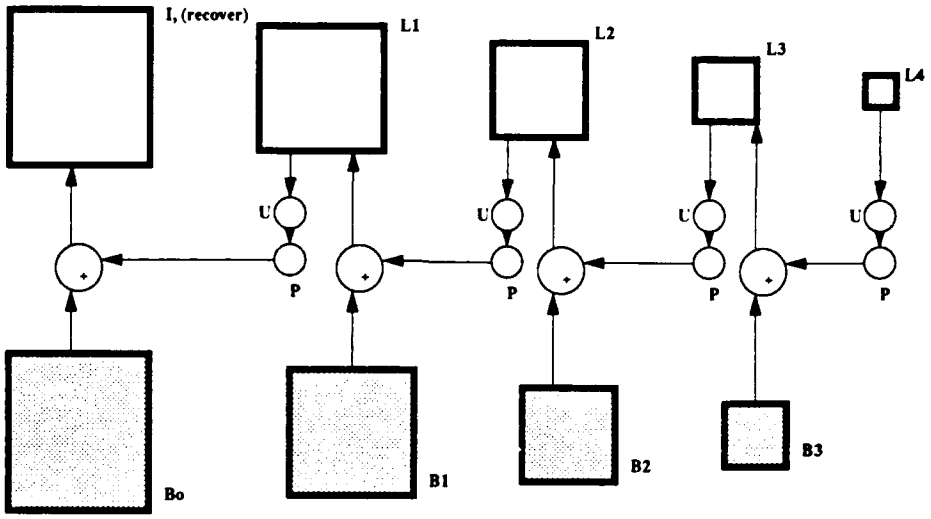

FIGURE 3. Course to fine synthesis process

Figure 3 shows the image synthesis process which recovers either the original or improved version from the bandpass pyramid, B_0 --> B_{k-2}, and the smallest lowpass image L_{k-1} which are shown as darkened images. The reconstruction process starts with up-sizing, U(.), L_{k-1} and then interpolating the larger version. The next step is to add in the smallest of the bandpass pyramids, Bk-2 to obtain Lk-2. These steps are then repeated all the way up the pyramid until reconstruction is completed.

$$L_i = (B_i + P(U(L_{i+1}))) \text{ for all i, i=1-->N-2, and}$$

$$I = L_0 = B_0 + P(U(L_1))$$

This method of reconstruction, course to fine synthesis, nicely lends itself to a controlled synthesis process. For small 2-D images this is not a real concern, but when 3-D imaging is attempted the controlled synthesis could likely save the user considerable time and effort when examining say 3-D medical data. Controlled synthesis works in the following manner: first at the coarse resolution, volumes of interest are identified; as the synthesis progresses only those regions identified as interesting are fully synthesized; the end result being that the user now has a 3-D image with the sections deemed interesting at the finest

resolution and the areas not of concern are 'roughed out.' This would be ideally suited for the progressive transmission of 3-D image data.

2.2 Burt's Laplacian Pyramid

The Laplacian pyramid proposed by Burt and Adelson in 1983 uses the identical construction process discussed above. Their goal was to use the pyramid structure as a method for image compression and they did get some decent results, over 5 to 1 on USC.girl with only 0.88% mean squared error and nearly 11 to 1 on the 'Walter' image with only 0.43% mean squared error. To achieve this result they further reduced the entropy in by quantizing the pixel values in each level of the pyramid images used in reconstruction. I used their 'best' pyramid as a benchmark for my nonlinear pyramids so a short overview of their work is required.

As previously stated, the general image pyramid construction procedures are identical to the process discussed earlier. In fact the only difference between their 'best' pyramid and all of my nonlinear pyramids is the filter, F(.), applied to the image to generate the lowpass pyramid. All the other operations are identical.

The filter they chose to use is a 5 x 5 weighted average which attempts to minimize the error after a level is down-sized, up-sized, and then interpolated. One dimensionally, Figure 4 shows a graphic representation at the pixel level of the Gaussian pyramid generation process.

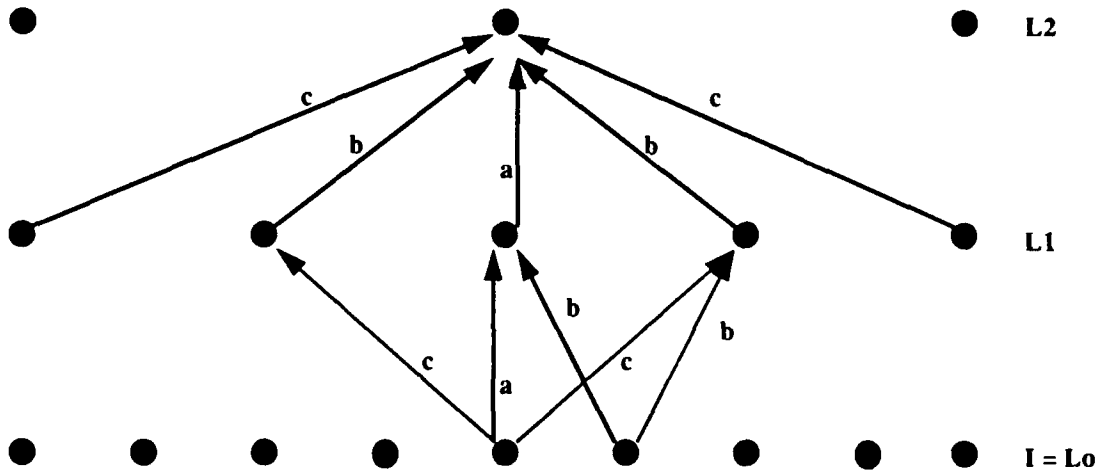

FIGURE 4. 1-D graphic representation of Gaussian pyramid generation process

The parameters 'a', 'b', and 'c' are the one dimensional weighting values used in the 5 x 5 convolution. The mask 'w(m,n)' is designed to be separable and symmetric with respect to the center of the mask, m=0 and n=0. Their second stipulation was that every pixel at a given level must contribute the same total weight (1/4) to the next level's nodes, equal

contribution. This is satisfied when $a + 2c = 2b$. It should also be noted that the mask is normalized to 1. These constraints are satisfied when: $(w(0) = a)$, $(w(1) = w(-1) = 1/4)$, $(w(2) = w(-2) = 1/4 - a/2)$. So the only option in their design is what value for a will be used. When a=0.4, this turns out to be, after repetitive application, a reasonable approximation to a Gaussian function and as the value for a decreases the function becomes more flat. The best results they obtained was when the value for a was chosen to be 0.6. The function is actually trimodel in nature when $a > 0.5$.

2.3 Nonlinear Pyramids

Now that the general pyramid's and Burt's Laplacian pyramid's construction procedures have been discussed, we turn to the nonlinear pyramids. The construction procedures for the nonlinear pyramids are no different than the general construction techniques. The only difference again being the filter used to generate the lowpass pyramid.

2.3.1 Basic Morphological Operators

For the benefit of any who actually read this report and may not be familiar with the basic morphological operations, I will give a very brief overview of the main operators used. There are two simple operators which define most morphological operators or filters, erosion and dilation. Let **B** represent the structuring element and **I** be the image. The gray-scale dilation and erosion of image **I** by structuring element **B** can then be defined as:

$$\text{Dilation} = I \oplus B = \text{MAX}_{x,y \in B}(I_{i+x,j+y} + B_{x,y}) \quad (\text{forall } (ij \in I))$$

$$\text{Erosion} = I \ominus B = \text{MIN}_{x,y \in B}(I_{i+x,j+y} - B_{x,y}) \quad (\text{forall } (ij \in I))$$

These definitions are assuming the structuring element is symmetric with respect to the center of **B** and **B**'s center is defined to be x = 0 and y = 0. For simplicity, the erosion by structuring element **B** will be written as $E_{\mathbf{B}}(.)$ and dilation as $D_{\mathbf{B}}(.)$. By combining these two operators, other morphological operators can be defined. The opening operator is simply an erosion followed by a dilation and closing is a dilation followed by an erosion.

$$\text{Opening} = O_B(I) = D_B(E_B(I))$$

$$\text{Closing} = C_B(I) = E_B(D_B(I))$$

2.3.2 Morphological Filters

Any one or combination of the above morphological operations can be used as the filter to generate the decomposition pyramid. The most commonly used filters are the opening and closing operations, though interesting results can be obtained by combining the basic operations in a different manner. When a linear type filter is used, the bandpass images generally represent different portions of the frequency domain; but if a morphological filter is used, the bandpass images are representative of some portion of the pattern spectrum in the image.

3.0 Image Compression

3.1 General Compression Measures

Two general measures of image compressibility are entropy and variance. If pixel values in an image are assumed to be statistically independent, then the entropy of the image yields the minimum number of bits per pixel (bpp) required to exactly encode the image. It is possible to approach this value using current compression techniques. Our test images contain 8 bpp corresponding to gray scale values from 0 to 255. Thus in the worst case the entropy of a given image would be 8.0 bpp while in the best case, a single valued image, the entropy would be zero.

$$H = -\sum_{i=0}^{255} f(i) \times \log_2 f(i)$$

'H' represents the image's entropy which is calculated from the images histogram; 'f(i)' is the normalized frequency of each gray scale value. In Burt's paper, he calculated the entropy of 'USC.girl' to be 7.57 bpp; while using my own software, I determined the entropy for the original image to be 7.59 bpp. So my software was reasonably accurate and supported their result.

Another representative measure of the compressibility of an image is variance. For the best case of a single valued image, the variance will be zero; while for the worst case for a 512 x 512 image, the variance would be 16,256.31. It must kept in mind that these two measures are not directly related. For example, the worst case for variance is in fact an image with a very low entropy, 1.0 bpp. The variance is calculated to be the normalized sum of the pixel differences from the mean then squared.

$$s^2 = \left(\sum_{i,j} (X_{ij} - X_m)^2\right) \times \frac{1}{n-1}$$

The result from Burt's paper for the variance of *USC.girl* is 2736 and the result I obtained for the same image was 2793. Again, the result is not perfect but very close.

3.2 Pyramid Compression

The motivation for using the pyramid structure is supplied in Burt and Adelson's paper on Laplacian pyramids Due to the high correlation of neighboring pixels, it is redundant to store image data as a simple matrix of pixel values. So a format is needed where the pixels are decorrelated if high compression rates are to be achieved. In predictive coding, the image is viewed as a single stream of pixel values, raster format, and previous pixel values are used to estimate the next pixel. Then the error between the predicted and actual value is what is stored for that pixel. Because prediction depends upon previously decoded values, this is said to be causal in nature [Burt 532]. Non-causal prediction based on a neighborhood should yield better prediction and thus better compression results. Non-causal prediction does suffer because simple sequential coding is no longer possible; techniques such as image transforms which encode blocks are generally used.

One point about this type of pyramid structure that must be considered if image compression is to be attempted is the pyramid is oversampled. This is one reason why there has not been much effort put into using the pyramid structure for data compression. The zeroth bandpass level contains the same number of pixels as the original image and then you must add the rest of the required levels. If we assume the 2-D image to be square with single dimension size of n, then the zeroth level contains n^2 pixels, the first level then has $n^2/4$, and so on. If we then assume a worst case where there are the maximum number of pyramid levels, k, the total number of pixels is a geometric series approaching $4n^2/3$.

$$\sum_{i=0}^{k} n^2 \times \left(\frac{1}{4}\right)^i \Rightarrow \frac{4n^2}{3}$$

3.3 General Compression Results

Part of the original goal of this project was to compare the usefulness of nonlinear and linear image pyramids for image compression. Two general measures of compressibility, entropy and variance, were used to compare the two types of pyramids. It is already known that the linear type of pyramid can provide good image compression despite the fact that it is over complete by 1/3 [Yu 161]. In general, the image pyramid decorrelates the pixel values leading to a set of histograms as pictured in figure 5.

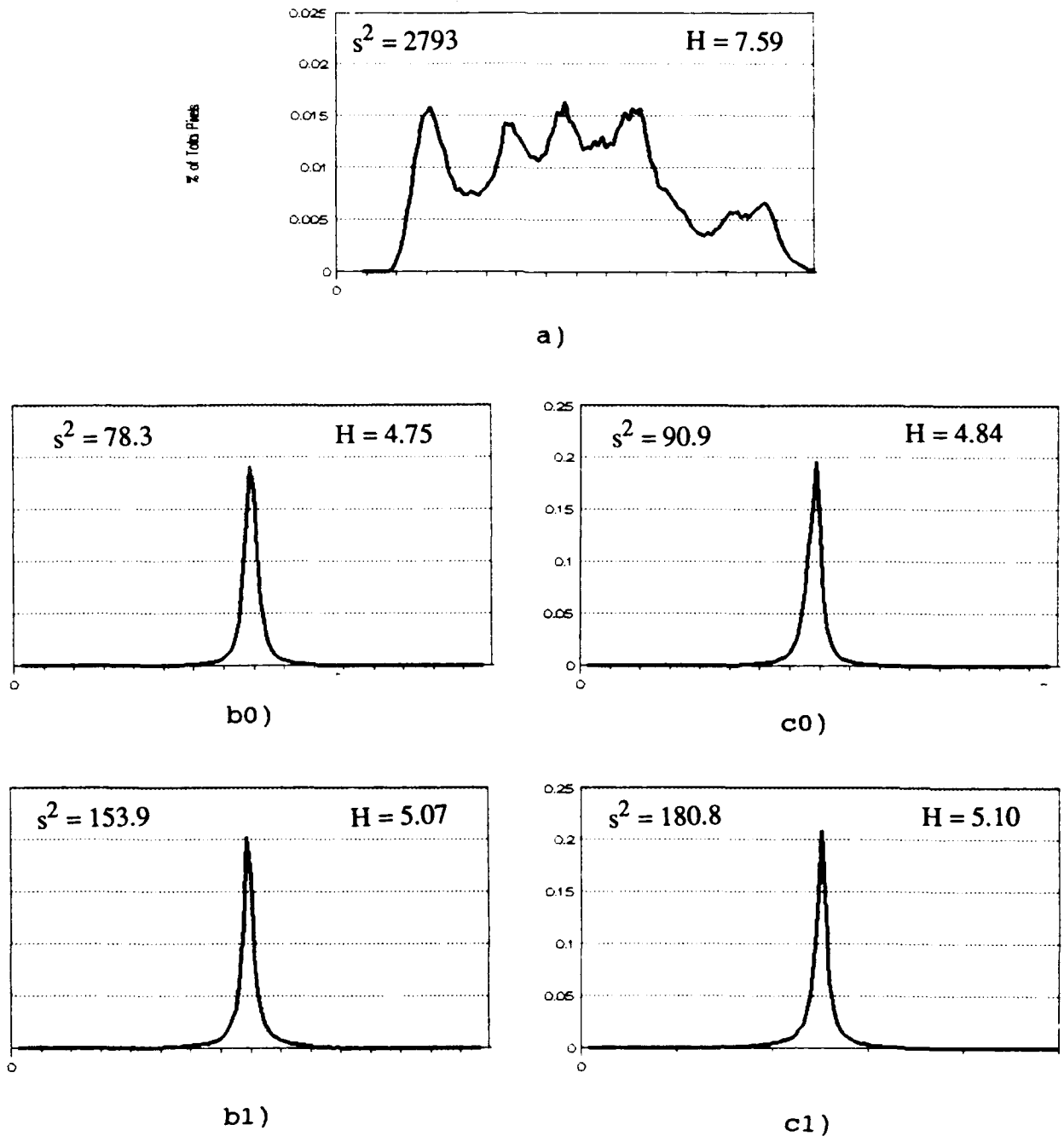

FIGURE 5. Histograms of (a) original USC.girl (b0) Burt.0 (c0) Close.0 (b1) Burt.1 (c1) Close.1

Figure 5.a is the original USC.girl and you can see that the image is somewhat random in its distribution and is thus very difficult to encode. Figure 5.b0 shows the highest level channel of the Laplacian pyramid, B_0, note that the distribution is now unimodal with entropy of 4.75 bpp and a variance of 78.3. The next level of theLaplacian pyramid, B1, is shown in figure 5.b1 and has an entropy of 5.07 bpp with a variance of 153.9. The best nonlinear pyramid performance, a closing by a level 3 x 3 cross filter, is shown in Figures 5.c0, B_0, and 5.c1, B_1. The variance of the top nonlinear pyramid level, B0, is 90.9 with a

corresponding entropy of 4.84 bpp; the next level of the nonlinear pyramid, B1, has an entropy of 5.10 bpp with a variance of 180.8. Note that the histograms of both the closing and linear pyramids are very similar in nature but have slightly different entropies and variances. It is difficult to see a difference at this scale but the differences become apperent when the histograms are zoomed in on. Due to the nature of the pyramid structure, the top level residual images must be highly compressible if the overall structure is to be compressible. The top levels contain most of the pixels in the image and thus only the top two levels were examined.

In an attempt to obtain better results than the Burt's Laplacian pyramid, we used three different structuring elements of varying heights. The two dimensional shapes of the structuring elements starting at a 3 x 3 cross and increasing in size on up to a diameter 5 disk. Constant valued and Gaussian shaped structuring elements were attempted with varying results. However, in general, the smaller the structuring element was, with height zero, the better it performed with respect to compressibility measures. Appendix A holds the results obtained for the different structuring elements and masks used. Several different images were also attempted, two cytological cell samples, and a slice from a 3-D image of a guinea pig cochlea, with the same general results.

These results show that the morphological pyramid can NOT provide consistently better image compression than Burt's Laplacian pyramid. The results do point out that the Opening and Closing pyramids always out perform the Erosion and Dilation types of pyramids. This should be obvious as the residue image represents the change in the original image, after *filtering*, *down_sizing*, up_sizing, and then interpolating, and the change is smaller when Opening or Closing is done.

4.0 Single Value Decomposition (SVD)

Burt stated in his Laplacian pyramid paper that the noncausal prediction techniques, such as pyramids, were best suited for image transforms such as SVD [Burt 532]. Thus, the next area of investigation was to see if the pyramid data structure, both linear and nonlinear, could use single value decomposition and recovery for data compression. This sections goal was to see if SVD could be used for compressing the different bands of the pyramid type of data structure. To be successful, the SVD of the pyramid must provide higher image compression than the SVD of the original image. A brief overview of the general SVD generation and reconstruction theory followed by the results obtained for this section will be given.

4.1 Transform Theory

Single value decomposition is a numerical technique used to diagonalize matrices in an attempt to achieve data compression [Andrews 425]. Deterministically, the SVD of an image is known to be the optimal transform for energy compaction [Jain 179], as the image is decomposed into a set of sorted eigenimages where they are sorted by the energy contained in each. Thus, by choosing only a small number of these eigenimages, it is pos-

sible to attain a high image compression ratio while minimizing the error due to compression loss. This will be later explained in more detail.

The original image, shown by the matrix [G], can be represented by a set of eigenvalues, or weights, and two orthonormal matrices [U] and [V].

$$[G] = [U][w][V]^T$$

'T' represents the transpose of the 'V' matrix. The matrix 'w' is the eigenvalues for the decomposed image and is actually all zero with the exception of the main diagonal which holds the sorted eigenvalues. There are several methods available to compute the 'U', 'w', and 'V' matrices including Symmetric QR, Hestenes Method, Golub-Kahan,.... The method used in this case was dictated by availability, 'Numerical Recipes in C' provides a decomposition routine based on the Golub method along with a reconstruction routine.

The reconstruction of the image from the 'U','w', and 'V' matrices is accomplished by multiplying each weight with the multiplication of corresponding vectors '$\mathbf{U}_i$' and '$\mathbf{V}_i^T$'; then adding up all the new matrices. Each 'i' represents what is known as an eigenimage. Due to the sorted nature of the eigenvalues, the first eigenimage contains the most information, then the second eigenimage, and so on.

$$G = \sum_{i=1}^{R} w_i \times \boldsymbol{U}_i \times \boldsymbol{V}_i^T$$

$$w_1 \geq w_2 \geq \ldots \geq w_R \geq 0$$

$$w_{R+1} = w_{R+2} = \ldots = w_n = 0$$

Depending on the specific image being decomposed, some of the weights may go to zero and thus be unnecessary for full image reconstruction. Some of the weights will also be very small relative to the highest energy weight, w_1. So it is then possible to use even fewer than 'R' weights, 'K', in the reconstruction and still recover the image with a minimal amount of error. Again, because of the energy packing property of SVD, this error is minimized in the following fashion.

$$G_k = \sum_{i=1}^{K} w_i \times U_i \times V_i^T$$

$$\|G - G_K\| = \sum_{i=K+1}^{R} w_i \times U_i \times V_i^T$$

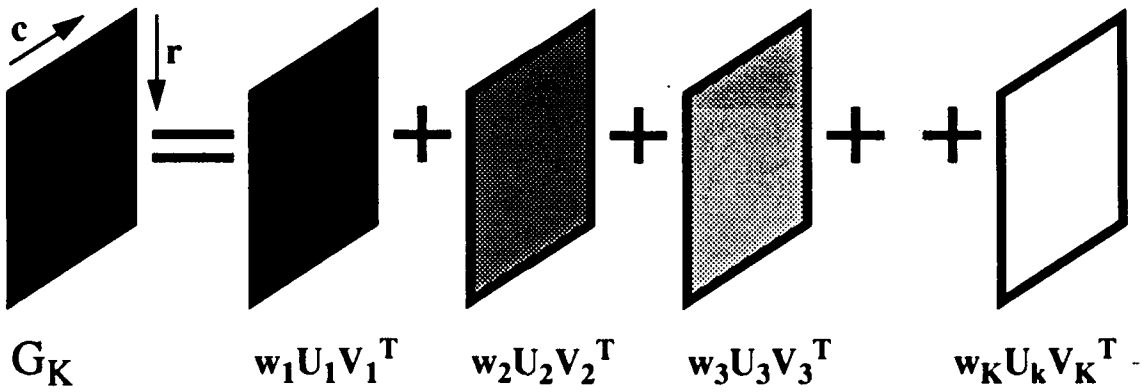

FIGURE 6. Graphical representation of eigenimages in reconstruction

Figure 6 shows graphically how the original image is reconstructed from the set of eigenimages. In the figure, the level of shading is representative of the amount of information present in image; the darker the image, the more information present.

4.2 Application Specifics

One of the major problems with using SVD is that the computation of the matrices 'U','w', and 'V' must all be computed for each individual image and this computation is $O(N^3)$ [Andrews1 426]. This is the worst computational efficiency of all the well known transform domains for image representations. Due to this problem and in the interest of not using all the CPU time, each image is divided into a set of smaller subblocks of dimension 64 x 64. Even with this subdivision of images, each pyramid, 512 x 512 original image, took well over 20 minutes to decompose.

If the comparison is to be accurate, the pyramid decompositions followed by SVD must give better results than the SVD of the original with the same number of values being stored. So, the relationship between the number of values for the original and for the pyramid decomposition must be derived. Let, 'S_b' represent the 1-D size of the subblock, 'n' be the 1-D size of the square original image, 'M_o' be the number of weights per subblock used in the recovery of the original image, and 'M_{pi}' represent the number of weights per

subblock to recover level 'i' of the pyramid. Each weight used corresponds to a single eigenimage and (2 * S_b + 1) values are required to store each eigenimage block.

$$(\frac{n}{Sb})^2 Mo(2Sb+1) \geq (2Sb+1)\left[(\frac{n}{Sb})^2 M0 + (\frac{n}{Sb}\frac{1}{2})^2 Mp1 + \ldots\right]$$

$$Mo \geq Mp0 + \frac{Mp1}{4} + \frac{Mp2}{16} + \frac{Mp3}{32} + \ldots$$

For simplicity sake, if it is assumed that Mp = Mp0 = Mp1 = Mp2 = - - - then the summation simply becomes a geometric series. The relationship between the number of weight used to recover the original must be at least 4/3 the number of weights used to recover the pyramid. This relationship is based on the recovered pyramid with the specified number of weights being of the same or better quality than the original recovered with its corresponding number of weights. If this relation holds true, then the pyramid SVD outperforms the

$$\frac{Mo}{Mp} \geq \frac{4}{3}$$

SVD of the original image.

The next question to be answered is what is better quality defined to be. For my work, I chose to use a combination of visual inspection and RMS error. RMS error is not always an ideal measure of image quality, but when used in conjunction with visual inspection, it can be a reasonable measure of image quality.

4.3 SVD Results

As previously stated, in order for the pyramid SVD to gain an advantage, it must have the same or better image quality than the SVD of the original image using 4/3 the number of weights as the pyramid SVD. For my actual comparison, see Figure 7, the 4/3 factor was not even taken into account; if it had been, the results would be even more obvious.

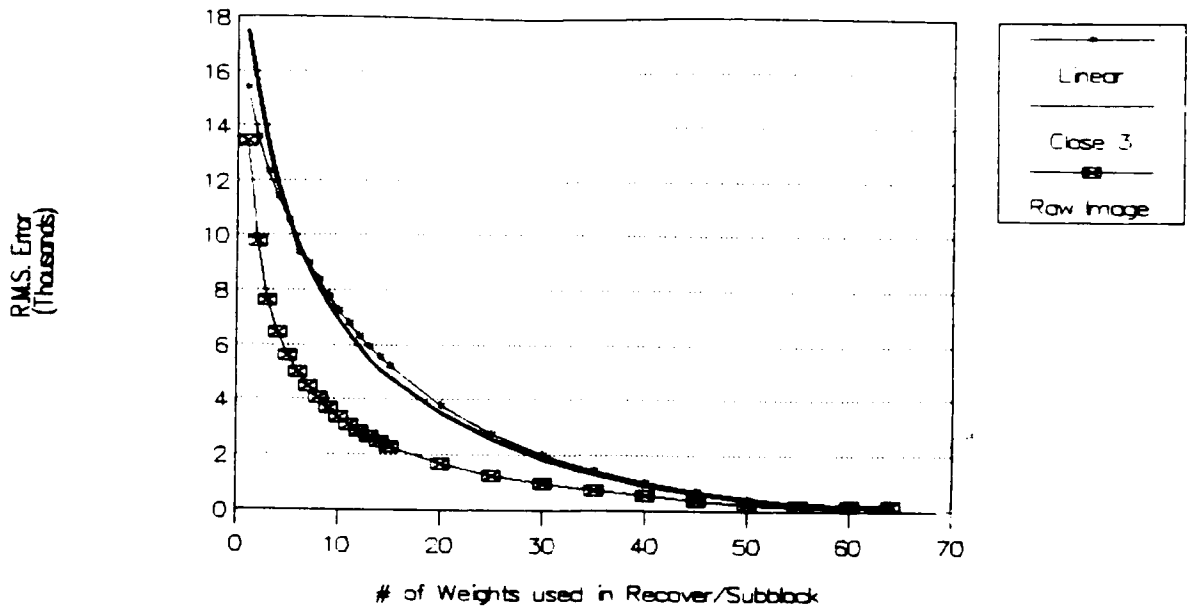

FIGURE 7. RMS error versus number of weights per subblock in reconstruction

As image quality is inversely related to the RMS error, Figure7 shows that the relationship between RMS error and the number of weights used for recovery is approximately exponential in nature, $(1 - ae^{-aM})$. However, the results were not what we were hoping to attain. The SVD of the original USC.girl consistently outperformed both the linear Laplacian pyramid SVD and the morphological pyramid SVD. There was not a noticeable difference between the linear and morphological SVDs. The morphological pyramid used for this example was a Closing by the 5 pixel diameter disk; it performed the best out of the nonlinear filters attempted.

Although the figure shows only RMS error, I did visually compare the SVDs at various points and it appeared that the RMS measure was providing an accurate representation of the image quality. So, I can come to no other conclusion than the combination of SVD and image pyramid combination is NOT an effective method for image compression.

Figure 8 shows the results for the reconstruction of *USC.girl*, 512 x 512, from (a) the raw image data and (b) a closing by 5 pixel diameter flat disk pyramid. The raw image was recovered using 15 weights persubblock while the bandpass pyramid was reconstructed using 12 weights per subblock, the 4/3 relationship holds true. It can easily be seen that the recovery from the raw data is much clearer and of better quality.

FIGURE 8. SVD Results: a) Raw image recovered, 15 weights b) Close (disk 5 pixel dia.) 12 weights

5.0 Pyramids as Image Processing Platforms

After the lack of success in the previous two attempts at using morphological pyramids, the application of the pyramids as an alternative image process platform was attempted. We wanted to check the possibility that morphological pyramid could outperform both the linear type of pyramid and the processing of the raw image data. To be successful at providing an alternative image processing platform, some special area or problem had to be better solved with a morphological type of pyramid.

First, the possible types of pyramid processing applications should be reviewed. These operations are not specific to morphological pyramids; instead, they are applicable to any type of image pyramid. Then, specifically what was accomplished in this area will be examined.

5.1 Potential Applications

5.1.1 Independent Band Processing

It is possible to process each band, or level, of the pyramid as a separate image in the same way that the raw image data is processed. However, because of the fact that each band is a portion of the pattern spectrum or frequency domain, different results will be achieved after the image is synthesized. Thus, a more powerful image processing platform is obtained.

Some examples of independent band processing include but are in no way limited to multiband noise coring and multiband interpolation. The noise coring operation can be accomplished by taking each high pass image, generated as the difference between the lowpass image and the previous lowpass image, and setting the pixel value to zero if it falls below a predetermined threshold value. Then, when the image is reconstructed from the bandpass and single lowpass image, the end result will be that much of the high frequency noise will be removed from the image. The loss of thudded information will be minimal in comparison to what would be lost if this operation were attempted at a single scale instead of a pyramid [Lee 16].

Multiband interpolation can be used to fill in missing information in an image. Though single scale interpolation operations exist, this method provides an alternative method to fill in the missing data. Figure 9 shows the method of interpolation using a morphological closing operation. Note that the missing data should be set to zero in both the lowpass decomposition pyramid and then in the bandpass pyramid.

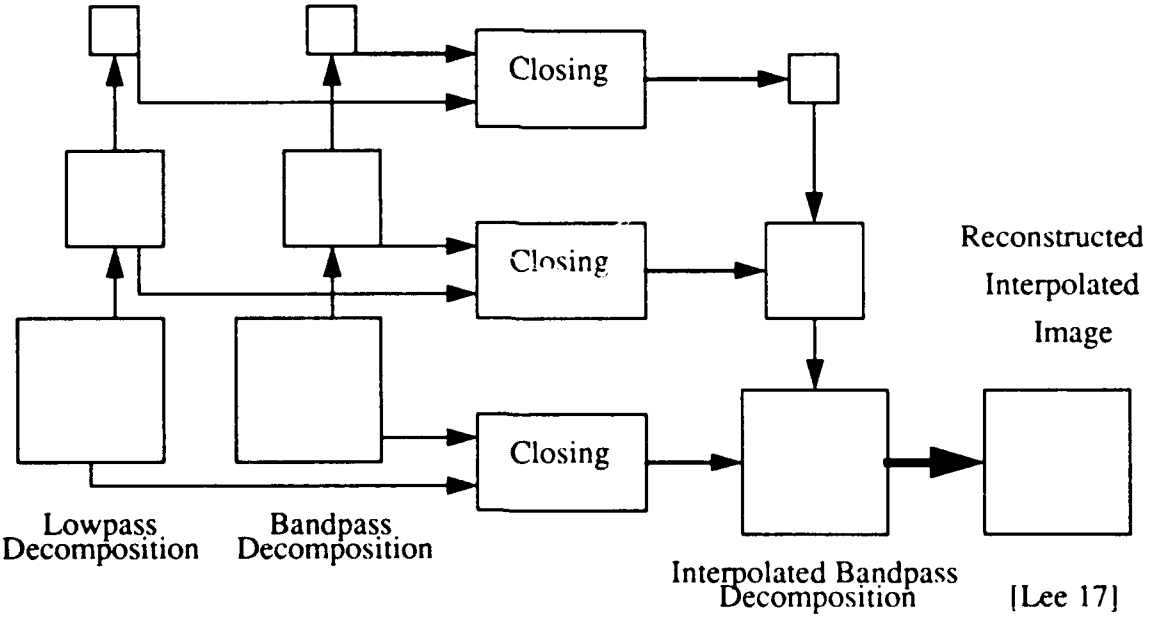

FIGURE 9. Multiband Interpolation

5.1.2 Cooperative Band Processing

Though the independent band processing does offer some unique possibilities, a more interesting application of image pyramids is cooperative band processing. Because of the pyramid property that spatial locality is maintained between pyramid levels, it is possible to perform operations between bands. This provides a method of processing unavailable when using raw image data.

Exceeding information can be removed from an image through the use of between band correlation which involves relating the data in a given band with other bands in the image pyramid. Two types of correlation operations that are available are a weakening correlation strategy and a strengthening correlation strategy.

The weakening correlation constrains the data intensity of a given band by taking the pixel by pixel minimum between the given bands and its adjacent bands. It is also possible to preprocess the adjacent bands with operations such as morphological dilation and then perform the weakening operation. For our application, the correlation was only done with one adjacent band B_{i+1}. In this case, the adjacent band, B_{i+1}, differs in size from B_i by a factor of two in each dimension, so it must first be up-sized then interpolated.

$$B_i^w = Min(B_i, Dilate(P(U(B_{i+1}))))$$

In this way, the resulting band image, B_i, is constrained to the dilated value of its adjacent band, B_{i+1}. This weakening operation can be applied iteratively from coarse (fine) resolution down to the fine (coarse) resolution [Lee 18].

The opposite of the weakening operation is enhancement. Image features can be enhanced, brightened, by performing the correlation between bands; but, instead of using a minimum operation as in weakening, a maximum operator is used and a type of weakening is first accomplished. This operation results certain features being enhanced with the size and degree of brightening being dependant upon the size and shape of the structuring element being used in the dilation and the structuring element used in the initial pyramid decomposition.

$$B_i^e = Max(B_i, Min(Dilate(B_i), (P(U(B_{i+1})))))$$

5.1.3 Selective Synthesis

By leaving out certain bands in the reconstruction it is possible to obtain different results. If the goal is to extract certain patterns below or above a certain size, then by using only

the bands containing the desired patterns with the given size in the reconstruction it is possible to obtain an image with the desired properties.

5.2 Specific Applications and Results

5.2.1 Strip Artifact

The first area that was investigated in hopes the morphological pyramid would prove superior was the removal of a structured noise, a strip artifact, from an image of a guinea pig cochlea, *Cochlea*, Figure 10.a. This type of artifact could possibly result from a digitizer with a neighborhood of photodetectors that were working incorrectly. The strip artifact introduced was five pixels wide with random bidirectional noise in the strip.Two types of linear pyramids with different frequency properties, a pair of morphological pyramids and classical noise removal techniques on the raw image data were attempted.

The morphological process used to remove the strip involves the use of two different types of nonlinear pyramids and weakening strategies. The first step was to generate an opening pyramid using a level 3 x 3 square structuring element. Then a weakening strategy using a dilation by the same structuring element of the smaller adjacent bands was used to remove part of the strip. The weakened bandpass pyramid was then reconstructed. The partially improved image was then again decomposed. This time, however, a closing pyramid based upon a flat disk, 5 pixels in diameter, was used. Then this bandpass pyramid was weakened using the smaller adjacent band dilated by the disk used in the decomposition. Finally, this bandpass pyramid was reconstructed resulting in the best removal of the strip, see Figure 10.b.

The two linear pyramids, each having a mask with different cutoff frequencies, used between band correlation to remove the strip. If the recovered images from both are examined, Figures 10.c and 10.d, it can be seen that both had partial success in removing the strip but failed to completely remove it. Even in the black area outside of the cochlea, the strip is still partially visible.

The classical attempt used was a repetitive convolution with a Gaussian type of lowpass mask. As can be seen in Figure 10.e, this attempt also failed to successfully remove the cochlea and predictably fared worse than the linear pyramids.

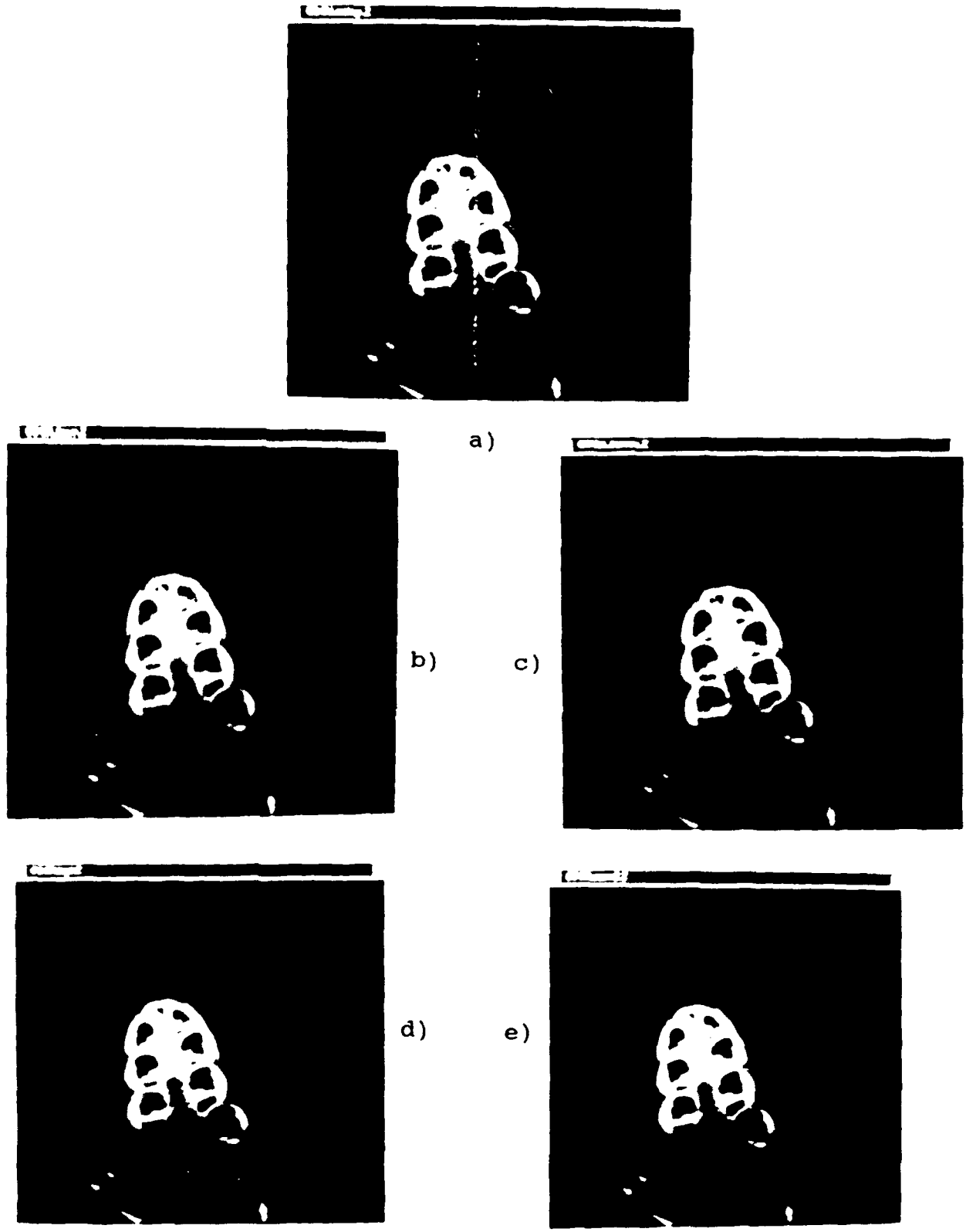

FIGURE 10. Guinea pig cochlea a) original b) Morphological c,d) Linear e) Raw Processing

5.2.2 Unstructured White Noise

After the morphological pyramid proved successful in removing the strip artifact with better results than the competing methods, the next logical step was to attempt the removal of noise from the entire image. Unstructured white noise, approximating speckle noise, was introduced to the *USC.girl* image resulting in figure 11.a. The next step was to attempt to remove the noise using a variety of techniques.

An opening pyramid based on a flat 3 x 3 cross structuring element was generated. This isolated the noise in the B_0 image on one side of zero. The next step was to perform a weakening of this band with a dilation by a flat 3 x 3 square of its adjacent band, B_1, to remove the noise. The image was then reconstructed using the normal process resulting in the removal of the noise with a minimal loss of image quality, see Figure 11.b.

For the two linear pyramids, each filter having a different cutoff frequency, I again attempted to use between band correlation to remove the noise. The weakening strategy had to be iteratively applied down the image pyramid as the noise was not isolated in one band. This resulted in greater error in the reconstructed image as the correlation operation does introduce some error. Figure 11.c shows the result of the first linear pyramid attempt and it fails to completely remove the noise. The other linear pyramid attempt does succeed in removing the noise but as Figure 11.d shows, a large amount of error was introduced.

The results would be incomplete if an attempt was not made to remove the same noise with classical techniques on the raw image data. Thus a repetitive covolution with a Gaussian lowpass mask was attempted. The results, Figure 11.e, clearly showed that the morphological pyramid to be superior.

Some might argue that a simple opening operation on the raw image data would be much simpler and as effective as the morphological pyramid. However, this is not true. If only opening is done then some information will be lost as more than just the noise is removed. When the pyramid is used, the noise is removed but the other information is added back into the image during the reconstruction process. Thus resulting in better performance.

As the *USC.girl* results show, the morphological pyramid proved superior to all other types of operations attempted. Each image has its corresponding RMS error which again shows the morphological approach to be the best for this application

a)
RMS =
8816.4

b) RMS = 2299.1

c) RMS = 3772.8

d) RMS = 3672.8

e) RMS = 4561.5

FIGURE 11. USC.girl a) speckle noise b) Morphological Pyramid c,d) Linear pyramid e) Convolution

6.0 Conclusions

This research was initiated to find some application in which morphological pyramid decompositions would provide the proper data representation, resulting in better results than current linear practices. Three different image processing applications of morphological pyramids were investigated: general image compression, image pyramid transforms, and pyramid processing.

For the general image compression measures, a comparison between Burt's Laplacian pyramid and a variety of morphological pyramids showed that the nonlinear pyramids could not outperform the linear decomposition. Though the linear pyramid consistently outperformed the nonlinear pyramids, some of the morphological pyramids did provide close results.

In the area of using the single value decompositions of image pyramids for compression, no success was found. The SVDs of the raw image data, a linear pyramid, and several morphological pyramids were examined. However, t he decomposition of the raw image da· proved far superior, with respect to lossy compression error, to all the pyramids attempted.

The final area of investigation was whether the nonlinear pyramids could provide a powerful image processing platform that could compete with both linear pyramids and classical raw image processing techniques. The results in t'· ·ection clearly showed the morphological pyramids could in fact be used for noise and artifact removal with much more success than classical raw image proce.sing and linear pyramid processing.

Our original goal of demonstrating a specific application where the nonlinear image pyramids could provide an equal or better method of data representation was accomplished. The nonlinear pyramid approach proved to be superior to other linear practices in the area of noise and artifact removal. No other method attempted could even come close to the results that were obtained in this area.

7.0 References

[1] H. C. Andrews and C. L. Patterson, "Single Value Decomposition (SVD) Image Coding." *IEEE Transactions on Communications*, April 1976.

[2] H. C. Andrews and C. L. Patterson, "Single Value Decompositions and Digital Image Processing." *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 24: 26 - 36, February 1976.

[3] P. J. Burt and E. Adelson. "The Laplacian Pyramid as Compact Image Code." *IEEE Transactions on Communication*, 31: 532 - 540, 1983.

[4] A. K. Jain. ***Fundamentals of Digital Image Processing***. Prentice-Hall: New Jersey, 1990.

[5] E. W. Kelm, J. N. Hwang, J. S. J. Lee, "Image Enhancement via Multiresolution Decomposition and Synthesis." Submitted to the *1992 ICASSP Proceedings*, San Francisco, March 1992.

[6] J. S. J. Lee, J. N. Hwang, and E. W. Kelm, "Image Analysis by Information Decomposition and Synthesis." To appear in the *25th Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, November 1991.

[7] Y. Lu am R. C. Vogt, "Multiscale Analysis Based on Mathematical Morphology." *SPIE Proceedings on Image Algebra and Morphological Image Processing II*, July 1991.

[8] S. Ranganath, "Image Filtering Using Multiresolution Representations." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13: 426 - 439, May 1991.

[9] P. Salembier and L. Jaqueoud, "Adaptive Morphological Multiresolution Decomposition." *SPIE Proceedings on Image Algebra and Morphological Image Processing II*, July 1991.

[10] T.H. Yu and S.K. Mitra, "A Simple Image Analysis / Synthesis Technique and Its Application in Image Coding." *SPIE Image Processing Algorithms and Techniques*, 1244: 161 - 170, 1990.

A P P E N D I X A

General Compressibility Results

LINEAR DECOMPOSITIONS

Dec. Filter:	Mask 1	Mask 2	Mask 3	Mask 3
Interpolator:	Mask 1	Mask 1	Mask 1	Mask 3
cellvar=	32.0117	44.5177	34.1990	43.2040
.0 H =	4.3284	4.5722	4.3812	4.5550
cellvar=	76.2796	66.8135	66.6020	73.7932
.1 H =	4.7928	4.5987	4.6548	4.6908
c4 var=	24.8096	34.4897	26.3730	31.5390
.0 H =	4.2835	4.5041	4.3262	4.4799
c4 var=	66.8618	63.5676	60.2686	66.6528
.1 H =	4.9499	4.9371	4.8838	4.9896
usc var=	78.3273	104.1386	82.3671	99.0217
.0 H =	4.7546	4.9567	4.7957	4.9359
usc var=	153.9719	105.6801	126.9528	118.0370
.1 H =	5.0710	4.8589	4.9208	4.9321
cochvar=	4.6887	10.8926	5.5107	9.7000
.0 H =	1.5399	1.9044	1.6701	1.8382
cochvar=	28.0918	28.9245	23.9694	20.9248
.1 H =	2.1917	2.2210	2.1529	2.2421

Morphological Open/Close Filters

Image	Open3	Open2	Open1	Close3	Close2	Close1
cellvar=	47.0396	38.1042	35.6470	61.6138	41.5765	37.0612
.0 H =	4.6386	4.4636	4.4042	4.7192	4.5000	4.4249
cellvar=	99.5010	89.0100	85.8693	153.3447	95.1517	83.3942
.1 H =	4.8881	4.8467	4.8392	4.7300	4.7391	4.7355
c4 var=	41.6664	30.6186	28.1678	38.1413	29.0155	27.1932
.0 H =	4.6190	4.4209	4.3624	4.5672	4.3959	4.3482
c4 var=	87.9576	74.0757	70.0855	118.7665	79.9035	73.1989
.1 H =	5.0819	4.9797	4.9494	5.2101	5.0284	4.9931
usc var=	164.9645	112.7772	95.2758	133.1236	98.9000	90.9297
.0 H =	5.0339	4.8766	4.8177	5.0955	4.9064	4.8388
usc var=	166.1879	158.5428	164.8461	204.5522	184.0000	180.7500
.1 H =	4.9157	4.9410	5.0142	5.1418	5.0794	5.1039
cochvar=	18.5640	11.7570	10.8354	7.1715	5.0484	4.6463
.0 H =	1.9110	1.7712	1.7017	1.7702	1.6129	1.5544
cochvar=	51.2361	35.9893	31.8528	29.1795	28.0779	28.3954
.1 H =	1.9951	2.0476	2.0809	2.3862	2.2419	2.1987

Morphological Dilate/Erode Filters

	Dilate3	Dilate2	Dilate1	Erode3	Erode2	Erode1
cellvar=187.6823	96.5397	68.9147	212.8918	97.6275	65.7415	
.0 H =	5.0270	4.6489	4.4975	5.2518	4.7645	4.5787
cellvar=199.7612	173.7714	145.3631	594.2546	280.5948	187.1824	
.1 H =	4.8042	4.6962	4.6892	5.8619	5.4070	5.2124
c4 var=	119.2735	60.2589	44.0260	123.7478	62.6035	45.3907
.0 H =	5.1457	4.6350	4.4557	5.2082	4.7125	4.5190
c4 var=	237.6876	154.3970	118.9445	319.1453	173.4942	124.6590
.1 H =	5.5716	5.2226	5.0707	5.8722	5.3985	5.1913
usc var=	425.2051	236.6170	168.1150	378.5289	215.0799	152.9301
.0 H =	5.5109	5.0958	4.9309	5.4045	5.0341	4.8945
usc var=	630.7359	431.9302	338.9891	465.7352	298.1260	246.9651
.1 H =	5.6159	5.3137	5.2379	5.4312	5.1359	5.0850
cochvar=	79.8322	35.8852	24.3568	72.4660	30.4239	19.3019
.0 H =	2.7994	2.2774	2.0049	2.2363	1.9559	1.7705
cochvar=	219.6546	98.5768	66.3602	106.2257	82.7359	64.6844
.1 H =	3.6326	3.0366	2.7760	1.9650	1.8745	1.8651

IMAGE ORIGINALS - Statistics

Cell Mean = 197.9106 variance = 1094.2032 Entropy = 6.6181

c4 Mean = 97.1678 variance = 1926.4603 Entropy = 7.4296

usc Mean = 126.0848 variance = 2793.2044 Entropy = 7.5930

coch* Mean = 9.1590 variance = 520.2139 Entropy = 2.9618

* 512 x 480

MASKS -

Mask - 1

0	-.03	-.05	-.03	0
-.03	.12	.25	.12	-.03
-.05	.25	.60	.25	-.05
-.03	.12	.25	.12	-.03
0	-.03	-.05	-.03	0

Mask - 2

9	13.6	16	13.6	9
13.6	24	32	24	13.6
16	32	64	32	16
13.6	24	32	24	13.6
9	13.6	16	13.6	9

L3 - recover W/ L3

16	32	16
32	64	32
16	32	16

Structuring Elements -

*1

```
 *
***
 *
```

*2

```
***
***
***
```

*3

```
  ***
 *****
 *****
  ***
```

NOTE : All morphological pyramids use the LB Mask to interpolate